A New Method of Digital Sound Reconstruction

Dominik Mayrhofer[1], Manfred Kaltenbacher[2]

[1] *Institute of Mechanics and Mechatronics, 1060 Vienna, Austria, Email: dominik.mayrhofer@tuwien.ac.at*

[2] *Institute of Mechanics and Mechatronics, 1060 Vienna, Austria, Email: manfred.kaltenbacher@tuwien.ac.at* &

Institue of Fundamentals and Theory in Electrical Engineering, 8010 Graz, Austria, Email: manfred.kaltenbacher@tugraz.at

Abstract

Classical Digital Sound Reconstruction (DSR) has been shown to be able to reproduce audio signals with the help of superimposed sound pulses. Still, this method has never been proven to be useful considering commercial applications. We start with the underlying idea of DSR and propose a new method – Advanced Digital Sound Reconstruction (ADSR) – which paves the way for a sound generation principle most effective in the lower frequency range. Using a redirection mechanism as an addition to classical DSR it is possible to achieve a completely flat frequency response regarding the sound pressure in a channel. This redirection unit uses shutter gates in order to separate positive and negative sound pulses letting only those pass, which contribute positively to the desired reconstruction of the audio signal. Compared to the well-known linear scaling of the sound pressure with the frequency, this concept offers exceptional performance for the low- to mid-frequency range – in the region where classical excitation schemes tend to struggle. Starting with the basic concept a concise framework of this method will be presented. This includes analytical investigations as well as numerical and experimental verification.

Introduction

Digital Sound Reconstrucion (DSR) is a concept that aims to improve the sound pressure level (SPL) especially for small devices like smartphones. Although traditional Micro-Electro-Mechanical Systems (MEMS) speakers are able to create enough sound pressure in the mid- to high-frequency range for such devices, below 1 kHz no sufficient sound pressure can be generated. The idea behind DSR is to increase the sound pressure by overlapping individual sound pulses to generate a signal similar to the function of an digital-to-analog-converter (DAC) known from electronics. By overlapping short pulses – which inherently create a large sound pressure due to their higher frequency content – the audio signal can be successfully reconstructed [1]. From this point on, more focus was laid upon developing digital loudspeaker arrays in order to incorporate the basic idea of DSR [2, 3, 4]. Although it is possible to recreate the signal in general, no major breakthrough in the aduio industry regarding an alternative to the classical analog mode has been recorded. It can be shown that – under certain circumstances – the underlying concept is not able to increase the achievable sound pressure and the only possibility to achieve better results with DSR is to alter the speaker characteristics favouribly. It has to be noted that other possible reconstruction techniques using modulation have been investigated as well, however, the main problem of not having a superiour reconstruction technique stayed the same [5]. In this work a new method called Advanced Digital Sound Reconstruction (ADSR) is discussed. With this concept it is possible to overcome the limitations of standard DSR by introducing a redirection process via a shutter device. This unit is able to generate an overall higher SPL, thus creating a new alternative to classical analog speakers which is ideally suited for a MEMS application due to its inherent properties. Combined with a new actuator this method paves the way for a new generation of loudspeakers capable of producing a high sound pressure especially at low frequencies.

Digital Sound Reconstruction

The basic idea of DSR is to overlap sound pulses to generate a desired sound pressure signal. The amplitude of the sound pulses corresponds to the actual value of the signal which should be generated, much like the classic principle of discretization. The discretized sound pressure is then assigned to the corresponding number of individual speaker cells, so called speaklets [1, 6]. A visualization of this process can be seen in Fig. 1.

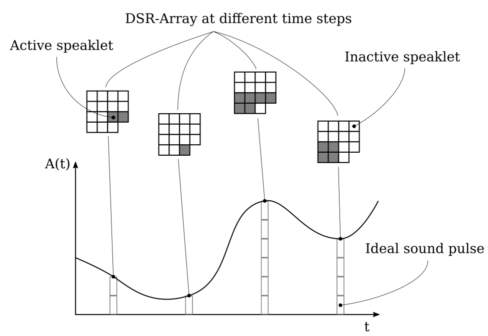

Figure 1: Visualization of the general idea.

It is important to note that each speaklet is either excited or not excited, but there is no variation of the amplitude of the excitation pulse, thus requiring some sort of array. Ideally this analog excitation pulse would give us an ideal positive or negative sound pulse as depicted in Fig. 1. Since this cannot be achieved with a classical actuator, ADSR was developed.

Advanced Digital Sound Reconstruction

The main idea to overcome the disadvantages of classical DSR is to redirect the unwanted sound waves at the scale of a single speaklet. In order to achieve this, the classical speaklet can be equipped with a shutter which manipulates the geometry in a favourable way. This combines the basic idea of DSR from the acoustic side with a redirection through a shutter mechanism which is only used in e.g. optics and can be seen as one way to realize a unit which we denote as an acoustic pump. Looking at the example of the positive half-period of a sine wave, the following operation mechanism can be observed:

- Opening the front- and closing the side-shutters.
- Exciting the membrane with a purely positive velocity signal and leaving the membrane at an elevated level, until the shutter position has been altered.
- Closing the front- and opening the side-shutters.
- Releasing the membrane back to its equilibrium state, which results in a negative sound peak, which is now redirected to the side and damped out.
- After returning to the equilibrium state, the process is started over again.

The described process now combines the advantage of generating a purely positive sound pulse by eliminating the problems of classical DSR by redirecting the unwanted part of the sound pulse. This procedure can be carried out with an exemplary device shown in Fig. 2.

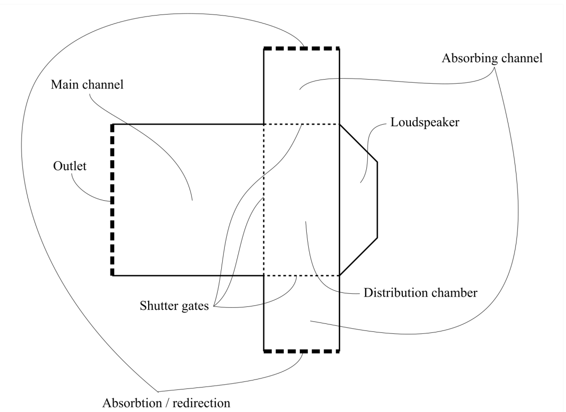

Figure 2: Exemplary actuator.

This device is the smallest unit which can be used to construct an array to be actually able to generate a music signal and will therefore be denoted as unit cell (UC). Considering the previously stated generalized solution it can be seen that, with the help of multiple speaklets equipped with shutter gates, it is possible to provide a continuous stream of pulses which corresponds to the basic principle of an acoustic pump and will be called a unit cell cluster (UCC). It has to be noted that we can either use a real DSR approach where we have an array of UCCs which are driven digitally, or, we can go back to a setting where we allow a variable amplitude of the pulses for a UCC, thus requiring only one such unit. Here, we have used the latter option for easier understanding.

Optimization

In order to evaluate the ideal excitation sequence as well as the optimal excitation signal a discrete iterative optimization process has been used, where the constrained optimization problem

$$\min_{x} \frac{1}{2} \left\| \boldsymbol{C} \cdot \boldsymbol{x} - \boldsymbol{d} \right\|_2^2 \quad \text{such that} \begin{cases} \boldsymbol{A} \cdot \boldsymbol{x} \leq \boldsymbol{b} \\ \boldsymbol{A}_{\text{eq}} \cdot \boldsymbol{x} = \boldsymbol{b}_{\text{eq}} \\ \boldsymbol{l}_{\text{b}} \leq \boldsymbol{x} \leq \boldsymbol{u}_{\text{b}}. \end{cases} \tag{1}$$

is solved. For example considering the optimization of the excitation signal in a global sense (the shape of the excitation signal stays constant and only the amplitude is varied), the Fourier-ansatz

$$g_{\text{exc}}(t) = c_0 + \sum_{i=1}^{N_{\text{H}}} a_i \cos\left(\frac{2\pi i t}{T_{\text{Dig}}}\right) + b_i \sin\left(\frac{2\pi i t}{T_{\text{Dig}}}\right). \tag{2}$$

can be used. The signal $g_{\text{exc}}(t)$ represents the excitation signal, N_{H} the number of harmonics, T_{Dig} the length of a purely positive or negative sound pulse and a_i, b_i and c_0 the Fourier-coefficients. Using a sinusiodal excitation sequence for a sinusiodal target signal results in the ansatz functions for the excitation signal for the whole target signal regarding the cosine-terms

$$\phi_i^c(t) = \left(\sum_{j=-1}^{N+1} \sin\left(\frac{2\pi j}{N}\right) \cos\left(\frac{2\pi i t}{T_{\text{Dig}^*}}\right) \big(H((j-1)T_{\text{Dig}}) - H((j+1)T_{\text{Dig}})\big) \right) \big(H(0) - H(T_{\text{audio}})\big) \tag{3}$$

as well as

$$\phi_i^s(t) = \left(\sum_{j=-1}^{N+1} \sin\left(\frac{2\pi j}{N}\right) \sin\left(\frac{2\pi i t}{T_{\text{Dig}^*}}\right) \big(H((j-1)T_{\text{Dig}}) - H((j+1)T_{\text{Dig}})\big) \right) \big(H(0) - H(T_{\text{audio}})\big) \tag{4}$$

for the sine-terms with N as the number of discrete points approximating the target signal and T^*_{Dig} the overall period of the pulse. The value T_{audio} denotes the period of the sinusiodal target frequency and H the Heaviside-function. For discrete values of a sinusoidal target signal stored in the target vector $\boldsymbol{d}$ as well as the ansatz given in (3) and (4) the system matrix $\boldsymbol{C}$ can be set up in order to solve for the design vector $\boldsymbol{x}$ of the form

$$\boldsymbol{x}^{\mathsf{T}} = \begin{bmatrix} b_0 & \dots & b_{N_{\text{H}}} & a_0 & \dots & a_{N_{\text{H}}} \end{bmatrix}. \tag{5}$$

The equality constraint matrix $\boldsymbol{A}_{\text{eq}}$, the equality constraint vector $\boldsymbol{b}_{\text{eq}}$, the inequality constraint matrix $\boldsymbol{A}$, the inequality constraint vector $\boldsymbol{b}$ as well as the upper and lower bounds $\boldsymbol{u}_b$ and $\boldsymbol{l}_b$ respectively can be used to control the behaviour of the target signal. By setting these constraints appropriately e.g. the start- and end-point or the integral value of the excitation signal can be controlled. The results of the optimization regarding the excitation sequence have shown that for a given excitation signal the ideal sequence for a sinusiodal target signal is again a sinusiodal, although a certain phase-shift might arise. For the ideal excitation signal one has to differentiate between the globally ideal excitation signal, where every pulse has a similar shape and only the amplitude is varied, or the case where every excitation signal is optimized with respect to the current section of the target signal. For the globally ideal excitation signal it has been found, that a triangular shape is ideal in terms of a high-frequency approximation. It has to be noted that, when dealing with a certain frequency band (e.g. 16 Hz to 16 kHz), the excitation signal leaves room for a minimal amount of fine tuning.

Analytical investigation

We define the time delay between two consecutive pulses as T_{Delay}. Furtermore we can define an overlap factor $o_f = T_{\text{Dig}}/T_{\text{Delay}}$ and a pause ratio $p_r = T^*_{\text{Dig}}/T_{\text{Dig}}$. By assuming $o_f = p_r = 2$ symmetry regarding the pulse widths is ensured which has been deemed as the most practical case. Additionally it enables us to derive a simple expression for the sound pressure amplitude $\hat{p}^{\text{d}}_{\text{a}}$ obtained with ADSR, which can be stated as

$$\hat{p}^{\text{d}}_{\text{a}} = \frac{\rho_0 c_0 s_{\max}}{2T_{\text{Dig}}} C_{\text{Area}} \tag{6}$$

for a channel as well as

$$\hat{p}^{\text{d}}_{\text{a,far}} = \frac{\rho_0 s_{\max} f_{\text{audio}} A}{2T_{\text{Dig}} z} \tag{7}$$

for the free field, where the latter one has been derived with the help of the Rayleigh integral. Hence, this is only valid for a large enough distance z between the array and the microphone. In these equations ρ_0 denotes the mean density of air, c_0 the speed of sound, $s_{\max}$ the maximum overall mechanical displacement and A the active area of the array. The factor C_{Area} takes the difference between the active area of the array and the actual area of the channel into account. If a non-pistophone like movement is prescribed, the displacement distribution has to be taken into consideration accordingly. For both cases a gain with respect to the analog mode can be calculated by

$$G = \frac{n}{4\pi}, \tag{8}$$

where n denotes the number of excitation pulses per target signal period. It has to be noted that although the gain is derived using the globally optimal excitation signal, multiple excitation signals have been tested where this relationship still gives a very good approximation.
Based on this derivation it can be clearly seen that we get a completely flat frequency response for a channel setup and only a linear scaling with the target frequency f_{audio} for the free field, which can also be seen in Fig. 3. Hence, for both cases a gain of 20 dB per decade can be obtained.

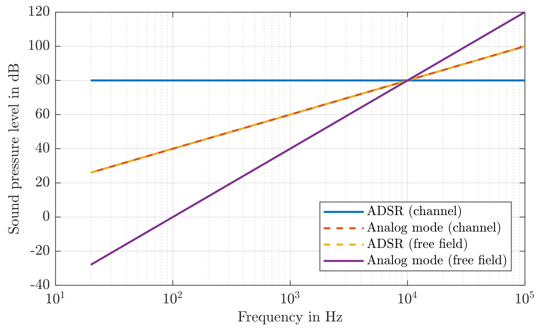

Figure 3: Comparison of the scaling laws for ADSR and the analog mode in a channel as well as the free field for an exemplary case.

Numerical computations

In the following, multiple acoustic simulations are presented in order to demonstrate the enormous benefit of ADSR in comparison to the analog mode. Therefore, we solve the weak form of the acoustic wave equation in the domain Ω in the potential formulation given by

$$\int_{\Omega} \psi' \frac{1}{c_0^2} \frac{\partial^2 \psi_{\text{a}}}{\partial t^2} \, \mathrm{d}\Omega + \int_{\Omega} \nabla\psi' \cdot \nabla\psi_{\text{a}} \, \mathrm{d}\Omega = \int_{\Gamma_{\text{n}}} \psi' \frac{\partial \boldsymbol{u}}{\partial t} \cdot \boldsymbol{n} \, \mathrm{d}\Gamma \, . \tag{9}$$

Here, ψ_{a} denotes the acoustic potential, ψ' the test function, $\partial \boldsymbol{u}/\partial t$ the prescribed velocity at the boundary Γ_{n} and $\boldsymbol{n}$ the corresponding normal vector. The computational setup seen in Fig. 4 is excited on one end by four equally large pistophones (speaklets) which correspond to the excitation area Γ_{n} and terminated with a perfectly matched layer on the other side. Regarding the excita-

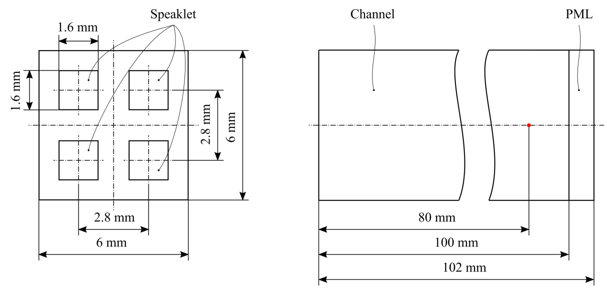

Figure 4: Sketch of the channel used for the acoustic simulations. The red dot marks the used microphone position for the evaluation (evaluation in the sweet spot).

tion for ADSR we assumed that only the desired positive or negative pulse is transmitted and therefore used for the prescribed velocity, which would require a perfect shutter mechanism. Furhtermore, it has to be noted that all simulations are matched regarding the maximum displacement of the individual cells in order to get compareable results. For the following simulations the parameters $s_{\max} = 600\,\text{nm}$ as well as $T_{\text{Dig}} = 62.5\,\mu\text{s}$ have been used. In order to investigate the derived scaling laws the analog mode as well as ADSR in combination with the globally ideal excitation signal have been compared for two frequncies, as seen in Fig. 5. Evidently, ADSR outperforms the classical analog mode amplitude

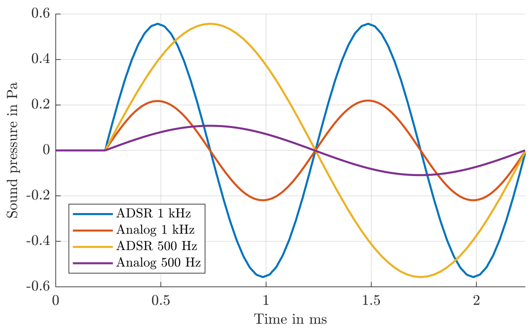

Figure 5: Comparison of ADSR and the analog mode for the frequencies 1 kHz and 500 Hz.

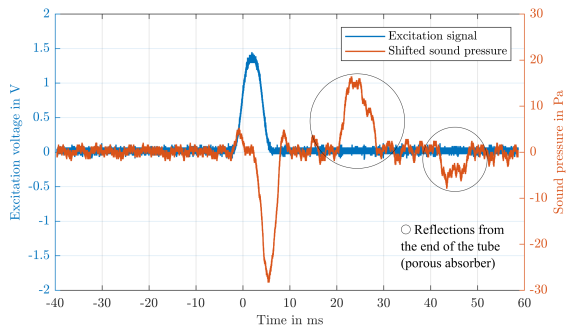

Figure 7: Proof of concept for the macroscopic unit cell based on a rotary shutter.

wise, but also the signal quality itself – in terms of Total Harmonic Distortion (THD) – is excellent. For the 1 kHz example the audible THD (16 Hz to 16 kHz) is 6.52 ppm, whereas the overall THD is 0.15 %. Regarding the 500 Hz example the THD is even lower, ranging from 2.82 ppm for the audible THD and 0.04 % for the overall THD. This clearly shows that if T_{Dig} is small enough, virtually no additional THD is induced in the audible range.

Measurements

In order to verify the principle of ADSR multiple measurement setups have been built, but only the most basic one will be presented. We present a single, macroscopic unit cell built based on a rotating shutter where we 3D-printed the shutter parts. A rotating core comprising the loudspeaker (Visaton FRWS 5 8 Ohm, broad band, 4W) as well as the rotating shutter part is driven by a stepper motor, where the rotating shutter part of the core is connected to the fixed shutter part in a casing, which is further connected to a measurement-channel, see Fig. 6.

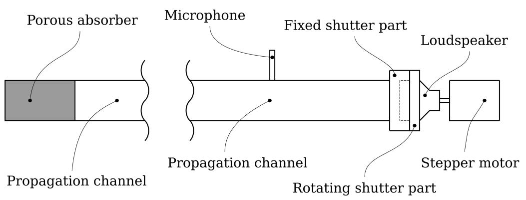

Figure 6: Sketch of the experimental setup for the proof of concept.

Using a smoothstep function for the excitation signal combined with the shutter mechanism has shown, that indeed a pureley positive or negative sound pulse can be achieved, see Fig. 7. In a further experiment, four unit cells have been combined in order to reconstruct a sinusiodal target signal, which has also been successful, although the real foundation is still the proof of concept for the operating principle of a single unit cell.

Conclusion

A new excitation concept for loudspeakers – called Advanced Digital Sound Reconstruction – was presented, which aims at a large performance gain regarding the achievable sound pressure in the low- to mid-frequency range especially for MEMS devices. Based on a redirection mechanism, a purely positive or negative sound pulse can be generated in a channel. This pulse is used to discretely approximate the desired audio signal in the sense of classical Digital Sound Reconstruction. Moreover, an optimization procedure was presented which has shown, that a triangular shape is ideal for a globally uniform excitation signal. Furthermore, we have presented analytical investigations which have been supported with numerical simulations as well as measurement setups. The provided framework demonstrates the enormous potential of ADSR: Regarding the sound pressure a completely flat frequency response in a channel and only a linear scaling with the frequency in the free field can be achieved.

References

[1] B. M. Diamond, J. J. Neumann, K. J. Gabriel: Digital sound reconstruction using arrays of CMOS-MEMS microspeakers. Fifteenth IEEE International Conference on Micro Electro Mechanical Systems (2002), ISSN: 1084-6999, 292-295

[2] A. Arevalo et al.: Towards a Digital Sound Reconstruction MEMS Device: Characterization of a Single PZT Based Piezoeletric Actuator. 10th IEEE International Conference on Nano/Micro Engineered and Molecular Systems (2015), 290-295

[3] S. Monkronthong, N. White and N.R. Harris: Multiple-level Digital Loudspeaker Array. Procedia Engineering, Vol. 87, 12 (2014), 404-407

[4] R. Dejaeger et al.: Development and Characterization of a Piezoelectrically Actuated MEMS Digital Loudspeaker. Procedia Engineering, Vol. 47, 09 (2012), 184-187

[5] D. Tumpold: Modeling and Optimization Methods of an Electrostatically driven MEMS Speaker. PhD thesis, TU-Wien (2014)

[6] J. J. Jr Neumann and K. J. Gabriel: CMOS-MEMS Acoustic Devices. CMOS-MEMS: Advanced Micro and Nanosystems (2008), Wiley-VCH, 193-224, ISBN: 9783527616718